Estatística II

Licenciatura em Gestão do Desporto
2.º Ano/2.º Semestre
2024/2025

Aulas Teórico-Práticas N.º 3 e 4 (Semana 2)

Docente: Elisabete Fernandes
E-mail: efernandes@iseg.ulisboa.pt

https://doity.com.br/estatistica-aplicada-a-nutricao

https://basiccode.com.br/produto/informatica-basica/

Conteúdos Programáticos

Aulas Teórico-Práticas (Semanas 1 a 3)	Aulas Teórico-Práticas (Semanas 4 a 7)	Aulas Teórico-Práticas (Semanas 7 a 9)	Aulas Teórico-Práticas (Semanas 10 a 13)
• **Capítulo 1**: Revisões e Distribuições de Amostragem	• **Capítulo 2**: Estimação	• **Capítulo 3**: Testes de Hipóteses	• **Capítulo 4**: Modelo de Regressão Linear Múltipla

Material didático: Exercícios do Livro Murteira et al (2015), Formulário e Tabelas Estatísticas

Bibliografia: B. Murteira, C. Silva Ribeiro, J. Andrade e Silva, C. Pimenta e F. Pimenta; *Introdução à Estatística*, 2ª ed., Escolar Editora, 2015.

Diferença de Médias Amostrais

1

Distribuições de Amostragem

Diferença de Médias Amostrais

Sejam $X_{11}, X_{12}, \dots, X_{1n_1}$ e $X_{21}, X_{22}, \dots, X_{2n_2}$ duas a. a. independentes, de dimensão n_1 e n_2 retiradas de duas populações com médias μ_1 e μ_2 e desvios padrão σ_1 e σ_2, respetivamente, e

$$\overline{X}_1 = \sum_{i=1}^{n_1} \frac{X_{1i}}{n_1} \text{ e } \overline{X}_2 = \sum_{i=1}^{n_2} \frac{X_{2i}}{n_2}.$$

ProbabilidadesEstatistica2019.pdf

Diferença de Médias: Variâncias Conhecidas

Se as populações forem Normais, sendo $X_1 \sim N(\mu_1; \sigma_1)$ e $X_2 \sim N(\mu_2; \sigma_2)$, com σ_1 e σ_2 conhecidos, então, pelo Teorema da aditividade da distribuição Normal, tem-se que:

$$Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0; 1).$$

Se as populações não forem Normais, mas as amostras forem de grande dimensão então, pelo corolário do T. L. C., vem $Z \overset{\circ}{\sim} N(0; 1)$, considerando o já anteriormente exposto em situação análoga.

ProbabilidadesEstatistica2019.pdf

Diferença de Médias: Variâncias Desconhecidas e Iguais

Se as populações forem Normais, sendo $X_1 \sim N(\mu_1; \sigma_1)$ e $X_2 \sim N(\mu_2; \sigma_2)$, com σ_1 e σ_2 desconhecidos mas iguais ($\sigma_1 = \sigma_2$), então demonstra-se que:

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}.$$

Se as populações não forem Normais, mas as amostras forem de grande dimensão então, por extensão do T. L. C., a expressão anterior segue aproximadamente uma $N(0; 1)$.

ProbabilidadesEstatistica2019.pdf

Diferença de Médias: Variâncias Desconhecidas e Diferentes

Se as populações forem Normais, sendo $X_1 \sim N(\mu_1; \sigma_1)$ e $X_2 \sim N(\mu_2; \sigma_2)$, com σ_1 e σ_2 desconhecidos e diferentes ($\sigma_1 \neq \sigma_2$), então, pela aproximação de Welch tem-se que (Murteira *et al.*, 2007):

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \mathrel{\dot\sim} t_v, \text{ onde } v = \left[\frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{S_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{S_2^2}{n_2}\right)^2}\right].$$

sendo $[r]$ a parte inteira de r, ou seja, arredonda-se por defeito o valor obtido.

Também neste caso, se as populações não forem Normais, mas as amostras forem de grande dimensão então, por extensão do T. L. C., a expressão anterior segue aproximadamente uma $N(0; 1)$, sendo válidas as observações anteriores.

ProbabilidadesEstatistica2019.pdf

Diferença de Médias Amostrais: Exercícios

2

Distribuições de Amostragem

Uma dada empresa farmacêutica lançou no mercado um novo medicamento, para dormir, que tem estado a ser utilizado nos hospitais. Constatou-se que os doentes não sujeitos a este medicamento em média dormiam 7,5 horas, com desvio padrão de 1,4 horas, ao passo que os doentes aos quais se administrou este medicamento dormiam em média 8 horas com desvio padrão de 2 horas.

Num determinado hospital observaram-se 31 doentes não sujeitos ao referido medicamento e 61 sob a referida medicação. Qual a probabilidade de os doentes do primeiro grupo observado dormirem em média mais do que os do segundo grupo? Assuma a normalidade das distribuições.

ProbabilidadesEstatistica2019.pdf

Exercício: Diferença de Médias Amostrais e Variâncias Conhecidas

Sejam:

- X_1 a v.a. que representa o número de horas de sono dos doentes não sujeitos ao medicamento,
- X_2 a v.a. que representa o número de horas de sono dos doentes sujeitos ao medicamento,

com

$$X_1 \sim N(\mu_1 = 7{,}5; \sigma_1 = 1{,}4) \text{ e } X_2 \sim N(\mu_2 = 8; \sigma_2 = 2).$$

$$\left.\begin{array}{l} X_1 \text{ e } X_2 \text{ dist. Normal} \\ \sigma_1\ (= 1{,}4) \text{ e } \sigma_2\ (= 2) \text{ conhecidos} \end{array}\right| \Rightarrow Z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0; 1).$$

$$P(\overline{X}_1 > \overline{X}_2) = P(\overline{X}_1 - \overline{X}_2 > 0) = 1 - P(\overline{X}_1 - \overline{X}_2 \leq 0) = 1 - P\left(Z \leq \frac{0 - (7{,}5 - 8)}{\sqrt{\frac{1{,}4^2}{31} + \frac{2^2}{61}}}\right) = 1 - \Phi(1{,}39)$$

$$= 1 - 0{,}9177 = 0{,}0823.$$

ProbabilidadesEstatistica2019.pdf

Uma dada empresa farmacêutica lançou no mercado um novo medicamento, para dormir, que tem estado a ser utilizado nos hospitais. Constatou-se que os doentes não sujeitos a este medicamento em média dormiam 7,5 horas, enquanto os doentes aos quais se administrou este medicamento dormiam em média 8 horas.

Num determinado hospital observaram-se n_1 doentes não sujeitos ao referido medicamento e n_2 sob a referida medicação tendo-se obtido, respetivamente, os seguintes desvios-padrão: 1,4 horas e 2 horas. Determine a probabilidade de os doentes do primeiro grupo dormirem em média menos do que os do segundo grupo, quando $n_1 = 20$ e $n_2 = 27$, quando se verifica a normalidade das distribuições e se considera:

a) A igualdade das variâncias populacionais.

b) A desigualdade das variâncias populacionais.

ProbabilidadesEstatistica2019.pdf

Exercício: Diferença de Médias Amostrais e Variâncias Desconhecidas e Iguais

Sejam:

- X_1 a v.a. que representa o número de horas de sono dos doentes não sujeitos ao medicamento,
- X_2 a v.a. que representa o número de horas de sono dos doentes sujeitos ao medicamento,

Com $X_1 \sim N(\mu_1 = 7{,}5; \sigma_1 =?)$ e $X_2 \sim N(\mu_2 = 8; \sigma_2 =?)$.

$n_1 = 20;\ \ s_1 = 20;\ \ n_2 = 27$ e $s_2 = 2.$

a)

X_1 e X_2 dist. Normal

σ_1 e σ_2 desconhecidos, mas iguais

$$\Rightarrow T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{n_1+n_2-2=45}.$$

$$P(\overline{X}_1 < \overline{X}_2) = P(\overline{X}_1 - \overline{X}_2 < 0) = P\left(T < \frac{0-(7{,}5-8)}{\sqrt{\frac{(20-1)1{,}4^2 + (27-1)2^2}{20+27-2}}\sqrt{\frac{1}{20}+\frac{1}{27}}}\right) = P(T < 0{,}957)$$

$$\approx 0{,}83.$$

ProbabilidadesEstatistica2019.pdf

Exercício: Diferença de Médias Amostrais e Variâncias Desconhecidas e Diferentes

b)

X_1 e X_2 dist. Normal
σ_1 e σ_2 desconhecidos e diferentes

$$\Rightarrow T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \overset{\circ}{\sim} t_{v=44},$$

pois

$$v = \left[\frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2} \right] = \left[\frac{\left(\frac{1,4^2}{20} + \frac{2^2}{27}\right)^2}{\frac{1}{20 - 1}\left(\frac{1,4^2}{20}\right)^2 + \frac{1}{27 - 1}\left(\frac{2^2}{27}\right)^2} \right] = [44,9] = 44.$$

$$P(\overline{X}_1 < \overline{X}_2) = P(\overline{X}_1 - \overline{X}_2 < 0) = P\left(T < \frac{0 - (7,5 - 8)}{\sqrt{\frac{1,4^2}{20} + \frac{2^2}{27}}}\right) = P(T < 1,008) \approx 0,84.$$

ProbabilidadesEstatistica2019.pdf

Murteira et al (2015)
Capítulo 6

41. Um centro emissor de cartões de segurança tem dois equipamentos de personalização, de funcionamento independente. O tempo de processamento, em segundos, para cada um deles tem comportamento normal com o mesmo tempo médio, sendo o desvio padrão do primeiro 10, e do segundo 15. Considerando amostras de 16 cartões personalizados de cada um dos equipamentos,

a) Calcule a probabilidade de a diferença entre as médias das duas amostras, em valor absoluto, ser superior a 5.

b) Qual a probabilidade do desvio padrão dos tempos de processamento dos cartões da amostra do primeiro equipamento ser superior ao da amostra do segundo equipamento?

Exercício 41 (a)

X_1 - v.a. tempo processamento do equipamento 1, em segundos

X_2 - v.a. tempo processamento do equipamento 2, em segundos

$X_1 \sim N(\mu_1, 10^2) \rightarrow$ Amostra casual : $m = 16$

$X_2 \sim N(\mu_2, 15^2) \rightarrow$ Amostra casual : $n = 16$

, onde $\mu_1 = \mu_2$

(a)

Quer-se : $P(|\bar{X}_1 - \bar{X}_2| > 5) = 1 - P(|\bar{X}_1 - \bar{X}_2| \leq 5) = 1 - P(-5 \leq \bar{X}_1 - \bar{X}_2 \leq 5)$

Sabe-se que : $Z = \dfrac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{m} + \dfrac{\sigma_2^2}{n}}} \sim N(0,1)$. Logo,

Exercício 41 (a)

$$1 - P\left(-5 \leq \bar{X}_1 - \bar{X}_2 \leq 5\right) = 1 - P\left(\frac{-5-0}{\sqrt{\frac{10^2}{16}+\frac{15^2}{16}}} \leq \frac{\bar{X}_1-\bar{X}_2-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}} \leq \frac{5-0}{\sqrt{\frac{10^2}{16}+\frac{15^2}{16}}}\right) =$$

$$= 1 - P\left(\underbrace{-\frac{5}{\sqrt{20.3125}}}_{(-1.11)} \leq Z \leq \underbrace{\frac{5}{\sqrt{20.3125}}}_{(1.11)}\right) = 1 - \left[\Phi(1.11) - \Phi(-1.11)\right] =$$

$$= 1 - \left[2\Phi(1.11) - 1\right] = 2 - 2\Phi(1.11) = 2 - 2 \times 0.8665 = 0.267$$

Variância Amostral

Distribuições de Amostragem

3

Variância Amostral

Seja $X_1, X_2, \ldots, X_n$, uma a. a. duma dada população com variância σ^2. A **variância amostral** é definida por:

Variância corrigida

$$S'^2 = \sum_{i=1}^{n} \frac{(X_i - \overline{X})^2}{n-1}.$$

Propriedades:

- $\mu_{S^2} = E(S^2) = \sigma^2.$
- Se a distribuição da população for Normal então $\sigma^2_{S^2} = Var(S^2) = \frac{2\sigma^4}{n-1}.$

Se a distribuição da população for Normal, com variância σ^2, então:

$$\chi^2 = \frac{(n-1)\, S'^2}{\sigma^2} \sim \chi^2_{n-1}.$$

$$Q = \frac{n\, s^2}{\sigma^2} \sim \chi^2(n-1)$$

ProbabilidadesEstatistica2019.pdf

Variância Amostral: Exercícios

Distribuições de Amostragem

4

Uma determinada empresa farmacêutica lançou no mercado um novo medicamento, para dormir, que tem estado a ser utilizado nos hospitais. Constatou-se que os doentes sujeitos a este medicamento em média dormiam 8 horas, sendo o desvio padrão de 2 horas, e que a distribuição do número de horas de sono podia ser considerada Normal.

Qual a probabilidade de, numa amostra aleatória de 31 doentes sujeitos ao referido medicamento:

a) A variância amostral ser superior a 5 horas?

b) A variância amostral ser inferior a 2,25 horas?

Supondo a variância corrigida

ProbabilidadesEstatistica2019.pdf

Exercício: Variância Amostral

Seja X a v.a. que representa o número de horas de sono dos doentes sujeitos ao medicamento, com $X \sim N(\mu = 8;\ \sigma = 2)$.

$$\left.\begin{array}{l} X \text{ dist. Normal} \\ n = 31 \end{array}\right| \Rightarrow \chi^2 = \frac{(n-1)\ S'^2}{\sigma^2} \quad S'^2 \quad \cdot 1 = 30 \cdot$$

Supondo a variância corrigida

a) $P(S^2 > 5) = 1 - P(S^2 \leq 5) = 1 - P\left(\chi^2 \leq \frac{(31-1) \times 5}{2^2}\right) = 1 - P(\chi^2 \leq 37{,}5) \approx 1 - 0{,}84 = 0{,}16.$

b) $P(S^2 < 2{,}25) = 1 - P\left(\chi^2 < \frac{(31-1) \times 2{,}25}{2^2}\right) = P(\chi^2 \leq 16{,}875) \approx 0{,}026.$

ProbabilidadesEstatistica2019.pdf

Murteira et al (2015)
Capítulo 6

39. De uma população normal de variância 64, tomou-se uma amostra de dimensão 3.

a) Qual a probabilidade de a variância da amostra exceder 78?

b) Responda à mesma pergunta para uma amostra de dimensão 16.

Supondo a variância não corrigida

Exercício 39 (a)

$X \sim N(\mu, 64) \longrightarrow$ Amostra : $n = 3$

(a)

Quer-se $P(S^2 > 78)$

Sabe-se que : $Q = \dfrac{nS^2}{\sigma^2} \sim \chi^2(n-1) = \chi^2(2)$, onde $n = 3$ e $\sigma^2 = 64$

Supondo a variância não corrigida

Logo,

$$P(S^2 > 78) = P\left(\frac{nS^2}{\sigma^2} > \frac{3 \times 78}{64}\right) = P(Q > 3.65625) \approx 0.1607$$

Exercício 39 (b)

Amostra : $n = 16$

$$Q = \frac{n s^2}{\sigma^2} \sim \chi^2(n-1) = \chi^2(15)$$

Supondo a variância não corrigida

Logo,

$$P(s^2 > 78) = P\left(\frac{n s^2}{\sigma^2} > \frac{16 \times 78}{64}\right) = P(Q > 19.5) \approx 0.192$$

Quocientes de Variâncias Amostrais

5

Distribuições de Amostragem

Quociente de Variâncias Amostrais

Sejam $X_{11}, X_{12}, \ldots, X_{1n_1}$ e $X_{21}, X_{22}, \ldots, X_{2n_2}$ duas a. a. independentes, de dimensão n_1 e n_2 retiradas de duas populações Normais, sendo $X_1 \sim N(\mu_1;\ \sigma_1)$ e $X_2 \sim N(\mu_2;\ \sigma_2)$, respetivamente, e

$$S_1^2 = \sum_{i=1}^{n_1} \frac{\left(X_{1i} - \overline{X}_1\right)^2}{n_1 - 1} \text{ e } S_2^2 = \sum_{i=1}^{n_2} \frac{\left(X_{2i} - \overline{X}_2\right)^2}{n_2 - 1},$$

então, por um teorema da distribuição F, tem-se que:

$$F = \frac{S_1^2}{S_2^2}\frac{\sigma_2^2}{\sigma_1^2} \sim F_{n_1-1;\, n_2-1}.$$

Supondo a variância corrigida

ProbabilidadesEstatistica2019.pdf

Quociente de Variâncias Amostrais: Exercícios

Distribuições de Amostragem

6

Uma determinada empresa farmacêutica lançou no mercado um novo medicamento, para dormir, que tem estado a ser utilizado nos hospitais. Constatou-se que os doentes não sujeitos a este medicamento em média dormiam 7,5 horas sendo o desvio padrão de 1,4 horas, enquanto os doentes aos quais se administrou este medicamento dormiam em média 8 horas sendo o desvio padrão de 2 horas.
Num determinado hospital observaram-se 31 doentes não sujeitos ao referido medicamento e 61 sujeitos ao medicamento. Qual a probabilidade da variância amostral do primeiro grupo ser inferior à do segundo grupo? Admita a normalidade dos dados.

Supondo a variância corrigida

ProbabilidadesEstatistica2019.pdf

Exercício: Quociente de Variâncias Amostrais

Sejam:

- X_1 a v.a. que representa o número de horas de sono dos doentes não sujeitos ao medicamento,
- X_2 a v.a. que representa o número de horas de sono dos doentes sujeitos ao medicamento,

com $X_1 \sim N(\mu_1 = 7{,}5; \sigma_1 = 1{,}4)$ e $X_2 \sim N(\mu_2 = 8; \sigma_2 = 2)$.

$$\left.\begin{array}{l} X_1 \text{ e } X_2 \text{ dist. Normal} \\ n_1 = 31 \text{ e } n_2 = 61 \end{array}\right| \Rightarrow F = \frac{S_1^2}{S_2^2}\frac{\sigma_2^2}{\sigma_1^2} \sim F_{n_1-1;\, n_2-1\, =\, 30;60}.$$

Supondo a variância corrigida

$$P(S_1^2 < S_2^2) = P\left(\frac{S_1^2}{S_2^2} < 1\right) = P\left(F < 1 \times \frac{2^2}{1,4^2}\right) = P(F < 2{,}04) \approx 0{,}99.$$

ProbabilidadesEstatistica2019.pdf

Murteira et al (2015)
Capítulo 6

41. Um centro emissor de cartões de segurança tem dois equipamentos de personalização, de funcionamento independente. O tempo de processamento, em segundos, para cada um deles tem comportamento normal com o mesmo tempo médio, sendo o desvio padrão do primeiro 10, e do segundo 15. Considerando amostras de 16 cartões personalizados de cada um dos equipamentos,

a) Calcule a probabilidade de a diferença entre as médias das duas amostras, em valor absoluto, ser superior a 5.

b) Qual a probabilidade do desvio padrão dos tempos de processamento dos cartões da amostra do primeiro equipamento ser superior ao da amostra do segundo equipamento?

Supondo a variância corrigida

Exercício 41 (b)

Então,

$$P(S_1'^2 > S_2'^2) = P\left(\frac{S_1'^2}{S_2'^2} > 1\right)$$

Supondo a variância corrigida

Sabe-se que: $F = \frac{S_1'^2}{S_2'^2} \cdot \frac{\sigma_2^2}{\sigma_1^2} \sim F(m-1, n-1)$, onde: $\sigma_1^2 = 10^2 = 100$; $\sigma_2^2 = 15^2 = 225$;

$m = 16$ e $n = 16$

Logo,

$$P\left(\frac{S_1'^2}{S_2'^2} > 1\right) = P\Bigg(\underbrace{\frac{S_1'^2}{S_2'^2} \cdot \frac{\sigma_2^2}{\sigma_1^2}}_{F(15,15)} > 1 \cdot \frac{225}{100}\Bigg) = P\left(F > 2.25\right) \simeq 0.064$$

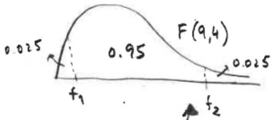
Murteira et al (2015)
Capítulo 6

45. De duas populações normais independentes, com variâncias iguais, foram extraídas duas amostras casuais com dimensões 10 e 5, respectivamente. Determine os valores tais que entre eles esteja, com 95% de probabilidade, a razão das variâncias corrigidas das amostras.

Exercício 45

$X_1 \sim N(\mu_1, \sigma_1^2) \longrightarrow$ Amostra casual : $m = 10$

$X_2 \sim N(\mu_2, \sigma_2^2) \longrightarrow$ Amostra casual : $n = 5$



$\sigma_1^2 = \sigma_2^2$

Quer-se encontrar os valores f_1 e f_2, tais que : $P\left(f_1 \le \frac{S_1'^2}{S_2'^2} \le f_2\right) = 0.95$

Sabe-se que $F = \frac{S_1'^2}{S_2'^2} \underbrace{\frac{\sigma_2^2}{\sigma_1^2}}_{1} \sim F(m-1, n-1) = F(9,4)$ . . Logo,

$$P\left(f_1 < \frac{S_1'^2}{S_2'^2} < f_2\right) = 0.95 \Rightarrow P\left(f_1 \times 1 < \frac{S_1'^2}{S_2'^2}\,\frac{\sigma_2^2}{\sigma_1^2} < f_2 \times 1\right) = 0.95 \Rightarrow P(f_1 < F < f_2) = 0.95$$

$$\Rightarrow \begin{cases} P(F < f_1) = 0.025 \\ P(F > f_2) = 0.025 \end{cases} \Rightarrow \begin{cases} P\left(\frac{1}{F} > \frac{1}{f_1}\right) = 0.025 \text{ , onde } \frac{1}{F} \sim F(4,9) \\ f_2 = 8.90 \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} \frac{1}{f_1} = 4.72 \\ \text{"} \end{cases} \Rightarrow \begin{cases} f_1 = \frac{1}{4.72} \\ \text{"} \end{cases} \Rightarrow \begin{cases} f_1 = 0.212 \\ f_2 = 8.90 \end{cases}$$

Obrigada!

Questões?